

用图型聚类法作气温变化趋势的概率预报

任荣彩 景 华 赵玉广 徐宝新

(河北省气象台, 石家庄 050021)

提 要

利用相似概率预报原理,通过图型聚类方法制作了河北省分冷暖季节、分辨站点的气温变化趋势的客观概率预报,1998年暖季开始试应用,暖季预报的布赖尔等评分结果表明,图型聚类方法是制作气温变化趋势概率预报的合理可行的方法。

关键词: 相似概率预报 气温变化趋势 图型聚类 布赖尔评分

引 言

近年来,我国部分省市相继建立了本地的降雨概率预报系统,1995年北京市气象局首先以概率形式向公众发布降水预报^[1],随后上海市、广州市也开始向公众发布降水概率预报^[2],在社会上引起了很大反响。1998年,河北省气象台完成了分季节、分量级、分观测站的降水概率客观预报系统^[3]。随着社会经济的发展,人类生活的决策方式日

益客观化、定量化,与降雨概率预报一样,其它气象要素的概率预报也愈来愈多地为公众及商业用户所需要。美国的 Somerville 和 Bean(1985)^[4]曾经利用概率密度函数建立的概率预报模式,预报天空云量、降水量及能见度等要素值大于或小于某值的概率;Scholefield(1977)^[5]利用图型相关法作各类云幕和能见度出现的概率的预报。

目前我国概率预报刚刚起步不久,我们针对公众和商业用户对气温显著变化的关心程度愈来愈高,而概率预报的形式更科学、更有利于客观定量决策,而引用了图型聚类方法,在相似概率预报原理基础上,制作最高(最低)气温显著升高、显著降低和基本不变的概率,其基本思想与文献[5]相似,都是通过对相关或相似图型的分(或归)类,来研究图型范围内天气现象与图型间的关系。不同的是,文献[5]中的图型是利用一种特有的图型设计技术事先设计出来,然后通过将历史图归到各图型,统计得到各型中预报对象的概率分布,预报则根据占优势的图型中的概率分布作出^[5];而本文概率预报所依据的图型,是通过对历史因子场进行系统聚类^[6]所归纳得到的,这种聚类方法是气象研究中采用较多的客观聚类方法,实践也证明,这种方法能够将图形属性相似的场归纳在一起,在被用于暴雨预报中的环流分型时,可以分辨各类暴雨型的基本特征^[7]。

本文合理选择因子场进行聚类分析,并基于各类图型作出河北省内11个市气温变化趋势的概率预报,而且通过合理的程序设计,研究了一套图型自动实时判别归类给出气温变化趋势概率值的业务预报系统。1998年暖季应用的布赖尔等评分结果说明,该系统有一定预报水平。说明图型聚类法是制作气温变化趋势概率预报的一种合理可行的方法。

1 资料处理及图型建立

1.1 资料处理

根据经验和历史(1993~1996年)最高、最低气温24小时变化情况,首先确定了气温变化的标准,即明显升高或明显降低或无明显变化,具体划分标准为:

明显升高:最高气温 $\Delta_{24}T_{\max} \geq 3.8^{\circ}\text{C}$;

最低气温 $\Delta_{24}T_{\min} \geq 3.3^{\circ}\text{C}$

明显降低:最高气温 $\Delta_{24}T_{\max} \leq -3.8^{\circ}\text{C}$;

最低气温 $\Delta_{24}T_{\min} \leq -3.3^{\circ}\text{C}$

无明显变化: $|\Delta_{24}T_{\max}| \leq 3.8^{\circ}\text{C}$;

$|\Delta_{24}T_{\min}| \leq 3.3^{\circ}\text{C}$

以河北省内11个市观测站为对象,建立暖季(5~10月份)逐日最高气温及最低气温的24小时变化趋势(明显升高取1;明显降低取0;无明显变化取-1)历史序列;冷季(11月份~次年4月份)逐日最低气温的24小时变化序列。

欧洲中期天气预报中心(EC)数值预报模式的预报能力已经为大家所公认,所以在选择与气温变化相关的因子场时,我们选择了EC的850hPa24小时变温预报场,截取资料的范围为:20~60°N、70~140°E,从而可相应地建立起逐日的因子场历史序列。

1.2 聚类图型建立

鉴于变温预报与气温变化的直接关系,以欧氏距离系数作为相似统计量,建立变温场的相似矩阵,然后采用平均串组法进行系统聚类分析^[6],其中距离系数计算公式为:

样本场 j 与样本场 k 之间的距离

$$d_{jk} = [(X_j - X_k)^T (X_j - X_k)]^{1/2}$$

类 G_j 与类 G_k 之间的距离

$$D_{JK} = \text{Average}(d_{jk}) \quad (\text{其中 } j \in G_j, k \in G_k)$$

聚类产生冷、暖季节的各类图型后,即可根据各图型所含样本,由逐日气温的24小时变化序列,分别统计出分冷、暖季节,系于各图型的省内11个观测站的24小时气温变化趋势的条件概率分布(以承德市站为例,见表1)。

表1 承德市站冷、暖季节气温变化趋势的概率分布(%)

类	冷季			暖季					
	最低气温			最高气温			最低气温		
	升	降	平	升	降	平	升	降	平
I	10	13	77	0	66	34	0	67	33
II	16	3	81	27	9	64	0	27	73
III	15	20	65	5	17	78	12	13	74
IV	7	7	86	15	8	77	14	11	75

2 气温变化趋势相似概率预报

由每日实时截取的EC 850hPa 24小时变温预报场,采用与聚类分析相同的统计量,

分别计算出实时场与各型样本平均场的距离系统 r_i , 选择最小值 R_k :

$R_k = \min_{1 \leq i \leq n} (r_i)$ (r_i 代表与 i 类的距离; n 为分型数), 将实时场归入第 k 类, 参照各型条件概率的分布, 从而可得到各观测站未来 24 小时气温变化趋势的概率预报值。

3 预报检验

专门用于检验概率预报的布赖尔 (Brier) 评分^[1] B 值越小, 说明预报概率值越接近于预报对象实况出现的频率, 即预报越准确, 但当预报对象的气候概率很小 (即小概率事件) 时, 因为 B 值 (代表预报概率与实况频率的差别大小) 很有可能与预报对象的气候概率值大小相当, 这时 B -评分即失去意义; 这时候可用技巧评分 B_s , 衡量概率预报准确率是否高于以预报对象的气候概率值作为预报值时的准确率, B_s 值大于零时则说明概率预报相对于气候概率有一定预报技巧, 否则即无技巧; 还可结合偏差评分 B_{ias} 值, 检验预报概率的倾向性, B_{ias} 大于零则说明预报概率相对于气候概率有预报偏多倾向, 相反有偏少倾向。

表 2 最高气温变化趋势的气候概率及预报评分结果

	无明显变化			明显升高			明显降低		
	R_0	B	B_s	R_0	B_s	B_{ias}	R_0	B_s	B_{ias}
张家口	0.77	0.18	+	0.11	+	-	0.12	+	+
承德	0.80	0.19	-	0.12	+	-	0.08	+	+
唐山	0.87	0.11	+	0.07	+	-	0.06	+	+
秦皇岛	0.83	0.13	+	0.08	-	-	0.09	+	+
廊坊	0.81	0.18	-	0.09	-	-	0.10	-	+
保定	0.79	0.16	+	0.09	-	-	0.12	+	-
沧州	0.83	0.13	+	0.08	+	-	0.09	+	-
衡水	0.82	0.13	+	0.08	-	-	0.10	+	+
石家庄	0.76	0.18	-	0.12	-	-	0.12	-	+
邢台	0.81	0.13	+	0.09	+	-	0.10	+	+
邯郸	0.83	0.14	+	0.08	+	-	0.09	+	+
平均	0.81	0.15	4.3	0.09	-7.6	-33	0.10	12.0	30.0

表 3 最低气温变化趋势的气候概率及预报评分结果

	无明显变化			明显升高			明显降低		
	R_0	B	B_s	R_0	B_s	B_{ias}	R_0	B_s	B_{ias}
张家口	0.80	0.18	+	0.11	+	-	0.09	+	+
承德	0.77	0.19	+	0.11	+	+	0.12	+	+
唐山	0.81	0.11	+	0.08	+	+	0.11	+	+
秦皇岛	0.90	0.13	+	0.04	+	+	0.06	+	+
廊坊	0.82	0.18	+	0.08	+	+	0.10	+	+
保定	0.88	0.16	+	0.05	+	-	0.07	+	+
沧州	0.84	0.13	+	0.08	-	-	0.08	+	+
衡水	0.84	0.13	+	0.07	+	+	0.09	+	+
石家庄	0.89	0.18	-	0.06	+	+	0.05	+	+
邢台	0.89	0.13	+	0.04	+	-	0.07	+	+
邯郸	0.90	0.14	+	0.04	+	+	0.06	+	+
平均	0.85	0.15	25.0	0.07	25.0	1.47	0.08	39.0	26.0

表 2 及表 3 给出了 1998 年暖季逐日预报的各站最高、最低气温变化趋势的评分结果及各温度趋势出现的气候概率 (对冷季预报的评分将在以后的工作中作出), 由表可见, 明显升高和明显降低趋势的气候概率 (R_0) 在 0.10 左右, 相对于无明显变化为小概率事件, 因此升高和降低趋势的概率预报评分以技巧 B_s 和偏差 B_{ias} 为主。由表可见, 最高、最低气温无明显变化的气候概率 (R_0) 在 0.8 左右, B 评分值均在 0.1~0.2 之间, 两值比较说明, 虽然个别站技巧分小于零, 但气温无明显变化的预报概率与实况频率的差别相对非常小, 亦即预报有相当高的水平。对于明显变化 (升高、降低) 趋势的概率预报的 B_s 技巧评分, 最低气温明显升高、明显降低和最高气温明显降低的 B_s 评分对于绝大多数观测站都大于零, 说明对于明显变化 (升高、降低) 趋势的概率预报, 除最高气温明显升高外都有一定预报技巧; 另外, 由表 2 和表 3 的对比还可以看出, 对最低气温明显变化的概率预报技巧好于最高气温; 无论最高还是最低, 明显降低趋势的概率预报技巧均好于明显升高, 这说明, 所选因子场对于各地市最低气温

变化的指示性(分辨力)较对最高气温的变化更强,对于各地气温明显降低趋势的指示性好于对气温明显升高趋势的指示性。这一方面说明,所选用因子场有一定局限性,由表1的概率分布也可以看出因子场对最高气温明显升高趋势的分辨力相对偏低;另一方面也可能说明,引起各地最高气温明显升高的原因较复杂,即除了形势分布特征外,还有地形等其它原因。另外,由表2及表3的 B_{ias} 分布可知,除最高气温明显升高趋势有预报偏少倾向外,对其它趋势的概率预报都有偏多倾向。这些都可以给日常气温预报中使用EC产品和概率预报产品提供客观参考。

4 结论

(1)选取合理的因子场,在相似概率预报原理基础上,用图型聚类方法制作的气温变化趋势的概率预报,有一定预报水平。气候概率(R_0)在0.8左右时,其B-评分值大都在0.1左右(表2、3)(明显升高和降低趋势的B评分在0.07~0.1之间,表略),相对比区域降雨概率预报的B-评分值小(降水气候概率仅0.2~0.3, B-评分值达0.1~0.2^[1]),亦即预报准确率更高。说明此方法是制作气温变化趋势概率预报的合理可行的方法。

(2)图型聚类方法的关键是所选取的因子场的代表性,即足够的样本及与预报对象

之间较高的相关性。

(3)若拥有更多的历史样本,可将图型聚类方法延伸为多因子场的组合相似聚类,从而有可能提高图型对各类气温变化趋势的分辨力,提高概率预报的评分水平。然而本文中由于资料样本的限制,采用的图型聚类方法尚属一种粗略的框架,还有待充实和改进,但却已经显示出了该方法在气温变化趋势概率预报方面的应用潜力。

致谢:我们的研究工作曾得到北京市气象局吴正华正研高工的指导,在此表示感谢!

参考文献

- 1 周兵等. 北京地区夏季降水概率预报业务应用研究. 气象, 1996, 22(1): 2~6.
- 2 谢定升等. 广州降水概率预报. 气象, 1998, 24(1): 43~46.
- 3 任荣彩等. 河北省降雨概率预报系统及其业务应用. 河北气象, 1999, 18(1).
- 4 Paul N. Somerville. et al., Probability Modeling of Weather Elements. 'Ninth Conference on Probability and Statistics in Atmospheric Sciences'. 1985: 173~175.
- 5 P. R. Scholefield. 用图形相关法作云量和能见度预报的概率预报模式. 第五届大气科学概率与统计会议文集(译文), 1977.
- 6 黄嘉佑等. 气象统计分析与预报方法. 北京: 气象出版社, 1990.
- 7 汤桂生等. 聚类分析在暴雨预报和环流形势分型中的应用. 气象, 1996, 22(8): 33~38.

Probability Forecast of Temperature Variation Tendencies with Pattern Cluster Technique

Ren Rongcai Jing Hua Zhao Yuguang Xu Baoxin

(Hebei Meteorological Observatory, Shijiazhuang 050021)

Abstract

Based on principles of similar probability forecast, the objective probability forecasts of temperature variation tendencies at stations were made separately for both warm and cold seasons by a pattern cluster technique. Brier and other scores of warm season show that the cluster technique is rational and feasible in temperature variation tendencies' probability forecast.

Key Words: similar probability forecasting temperature variation tendency pattern cluster Brier score